



# ALGORITMA *STEMMING* TEKS BAHASA BALANTAK BERBASIS ATURAN TATA BAHASA

## LAPORAN TUGAS AKHIR

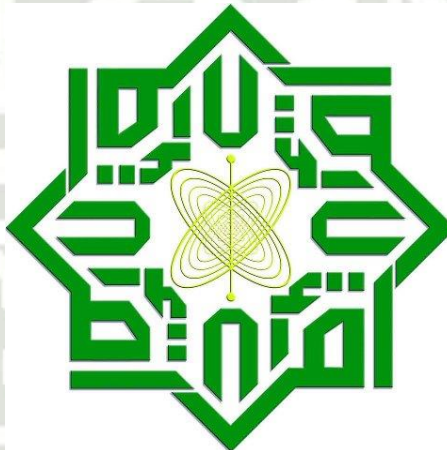
Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Teknik

Pada Jurusan Teknik Informatika

oleh:

**SURYA ATMAJA**

**11351101946**



**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU**

**PEKANBARU**

**2021**

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LEMBAR PERSETUJUAN

### ALGORITMA *STEMMING* TEKS PADA BAHASA BALANTAK BERBASIS ATURAN TATA BAHASA

#### TUGAS AKHIR

Oleh

**SURYA ATMAJA**

**11351101946**

Telah diperiksa dan disetujui sebagai laporan tugas akhir  
di Pekanbaru, pada tanggal 08 Januari 2021

Pembimbing I,

**Muhammad Fikry, S.T., M.Sc**  
**NIP. 19801018 200710 1 002**

Pembimbing II,

**Yusra, S.T., M.T**  
**NIP. 19840123 201503 2 001**

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.





## LEMBAR PENGESAHAN

### ALGORITMA *STEMMING* TEKS PADA BAHASA BALANTAK BERBASIS ATURAN TATA BAHASA

#### TUGAS AKHIR

Oleh

**SURYA ATMAJA**

11351101946

Telah dipertahankan di depan sidang dewan penguji  
sebagai salah satu syarat untuk memperoleh gelar sarjana Teknik Informatika  
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau  
di Pekanbaru, pada tanggal 08 Februari 2021

Pekanbaru, 08 Februari 2021

Mengesahkan,

Ketua Jurusan,

Dekan,



**Dr. Drs. Ahmad Darmawi, M.Ag.**  
NIP. 19660604 199203 1 004

**Dr. Elin Haecani, S.T., M.Kom.**  
NIP. 19810523 200710 2 003

#### DEWAN PENGUJI

Ketua : Iwan Iskandar, S.T., M.T  
Sekretaris : Muhammad Fikry, S.T., M.Sc  
Anggota I : Yusra ., MT  
Anggota II : Muhammad Affandes, S.T., M.T  
Anggota III : Suwanto Sanjaya, S.T., M.Kom.

iii



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

## LEMBAR HAK KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam.

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan didalam daftar pustaka.

Pekanbaru, 08 Januari 2021

Yang membuat pernyataan,

**SURYA ATMAJA**

**11351101946**

UIN SUSKA RIAU

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.





**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LEMBAR PERSEMBAHAN



Maka nikmat tuhanmu yang manakah yang kamu dustakan? (QS. Ar-rahman 13)

Assalamu'alaikum Warahmatullahi Wabarakatuh

Ku susun jari jemari ku diatas keyboard laptop ku sebagai pembuka kalimat persembahan. Diikuti dengan Bismillahirrahmanirrahim sebagai awal setiap memulai pekerjaan.

Sebuah langkah telah usai sudah, satu cita-cita telah tercapai. Sembah sujud serta puji dan syukurku pada-Mu Allah SWT. Tuhan semesta alam yang menciptakanku dengan bekal yang begitu teramat sempurna.

Taburan cinta, kasih sayang, rahmat dan hidayat-Mu telah memberikan ku kekuatan, kesehatan, semangat pantang menyerah dan memberkatiku dengan ilmu pengetahuan serta cinta yang pasti ada disetiap ummat-Mu. Atas karunia serta kemudahan yang Engkau berikan akhirnya tugas akhir ini dapat terselesaikan. Sholawat dan salam selalu ku limpahkan kepada Rasulullah Muhammad SAW.

Ku persembahkan tugas akhir ini untuk orang tercinta dan tersayang atas kasihnya yang berlimpah yaitu Ayahanda dan ibunda tercinta sebuah tulisan dari

didikan kalian yang ku aplikasikan dengan ketikan hingga menjadi barisan tulisan dengan beribu kesatuan, berjuta makna kehidupan, tidak bermaksud yang lain hanya ucapan TERIMA KASIH yang setulusnya tersirat dihati yang ingin ku sampaikan atas segala usaha dan jerih payah pengorbanan untuk anakmu selama ini. Hanya sebuah kado kecil yang dapat ku berikan dari bangku kuliahku yang memiliki sejuta makna, sejuta cerita, sejuta kenangan, pengorbanan, dan perjalanan untuk dapatkan masa depan yang ku inginkan atas restu dan dukungan yang kalian berikan.

Terimakasih untuk do'a - do'anya

Semoga tugas akhir ini bermanfaat bagi pembacanya Allahumma Aamiin



# ALGORITMA *STEMMING* TEKS BAHASA BALANTAK BERBASIS ATURAN TATA BAHASA

**SURYA ATMAJA**  
**11351101946**

Tanggal Sidang :

Jurusan Teknik Informatika  
Fakultas Sains dan Teknologi  
**Universitas Islam Negeri Sultan Syarif Kasim Riau**

## ABSTRAK

Bahasa Balantak merupakan salah satu bahasa daerah di Provinsi Sulawesi Tengah umumnya di daerah Balantak, Kabupaten Banggai dan sekitarnya. Pada penelitian ini aturan imbuhan terdapat dari buku struktur bahasa Balantak, kata dasar dari percakapan sehari-hari pada bahasa Balantak, serta buku struktur sastra lisan Balantak sebagai data uji. Dalam bahasa Balantak terdapat kata berimbuhan yang terbentuk dari awalan, akhiran dan perulangan. *Stemming* merupakan sebuah teknik ekstraksi sebuah kata berimbuhan dengan tujuan untuk mencari kata dasar dari kata yang berimbuhan dengan cara menghilangkan atau menghapus imbuhan yang terdapat pada kata dasar. Teknik *stemming* dibagi menjadi tiga tahapan yaitu tahapan mengumpulkan aturan yang terdapat pada bahasa Balantak, tahapan pengumpulan kata dasar yang akan menjadi kamus basis data dan tahapan penghapusan imbuhan. Perancangan yang dilakukan yaitu membuat *flowchart* dan *pseudocode* algoritma *stemming* teks bahasa Balantak. Berdasarkan hasil penelitian dan pengujian yang dilakukan menggunakan *white box*, algoritma *stemming* yang telah dibuat dengan menggunakan 310 kata uji diperoleh 2 hasil menggunakan kombinasi pertama 299 kata uji dengan hasil benar dengan total akurasi 96,45% dan 11 kata uji dengan hasil salah sebanyak 3,55%, kombinasi kedua 305 kata uji dengan hasil benar dengan total akurasi 98,39% dan 5 kata uji dengan hasil salah sebanyak 1,61%.

**Kata Kunci:** Algoritma, Bahasa Balantak, *Flowchart*, *Pseudocode*, *Stemming*, *White Box*



# **BALANTAK LANGUAGE STEMMING TEXT ALGORITHM BASED ON LANGUAGE RULE BASED**

**SURYA ATMAJA**  
**11351101946**

*Session Date:*

*Informatics Engineering Department  
Faculty of Science and Technology  
Sultan Syarif Kasim State Islamic University Riau*

## **ABSTRACT**

*Balantak language is one of the regional languages in Central Sulawesi Province generally in Balantak, Banggai Regency and surrounding areas. In this study, the rules were contained from balantak language structure book, the basic word of daily conversation in Balantak language, as well as balantak oral literary structure book as test data. In Balantak there are healing words that are formed from preces, sy endings and repetitions. Stemming is a technique of extracting a word with the aim of finding the basic word of the word that heals by eliminating or deleting the imbuhan contained in the base word. Stemming technique is divided into three stages, namely the stages of collecting rules contained in the Balantak language, the basic word collection stage that will become a database dictionary and the deletion stage. The design is to create flowchart and pseudocode algorithm stemming balantak text. Based on the results of research and testing conducted using white boxes, stemming algorithm that has been created by using 310 test words obtained 2 results using the first combination of 299 test words with correct results with total accuracy of 96.45% and 11 test words with incorrect results of 3.55%, the second combination of 305 test words with correct results with a total accuracy of 98.39% and 5 test words with incorrect results of 1.61%.*

**Keywords:** *Algorithm, Balantak Language, Flowchart, Pseudocode, Stemming, White Box*





**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## KATA PENGANTAR

*Assalammu'alaikum warahmatullahi wabarakatuh.*

Alhamdulillah, segala puji hanya bagi Allah SWT, karena dengan berkah limpahan rahmat dan karunia-Nya penulis dapat menyelesaikan Laporan Tugas Akhir yang berjudul **“ALGORITMA STEMMING TEKS PADA BAHASA BALANTAK BERBASIS ATURAN TATA BAHASA”**. Juga berkat dukungan orangtua dan keluarga besar yang selalu memberikan semangat dan dorongan untuk kelancaran tugas akhir penulis.

Laporan Tugas Akhir ini disusun sebagai salah satu syarat kelulusan dari Universitas Islam Negeri Sultan Syarif Kasim Riau jurusan Teknik Informatika. Selama melaksanakan tugas akhir ini, penulis banyak mendapatkan pengetahuan, bimbingan, dukungan, dan arahan dari semua pihak yang telah membantu hingga penulisan laporan ini dapat diselesaikan. Untuk itu, penulis ingin menyampaikan ucapan terimakasih kepada:

1. Bapak Prof. Dr. Suyitno, M.Ag, selaku PLT Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Drs. Ahmad Darmawi, M.Ag selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
3. Ibu Dr. Elin Haerani, S.T., M.Kom, selaku Ketua Program Studi Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Bapak Teddi Darmizal, S.T., M.Sc., selaku Pembimbing Akademik Penulis selama menjalani perkuliahan di Jurusan Teknik Informatika.
5. Bapak Muhammad Fikry, S.T., M.Sc., selaku Pembimbing I penulis dalam melakukan penelitian Tugas Akhir yang telah banyak memberikan ilmu, wawasan serta saran dan arahan sehingga penulis dapat menyelesaikan penyusunan dan penulisan Tugas Akhir dengan baik dan benar.
6. Ibu Yusra, S.T., M.T., selaku Pembimbing II penulis yang telah banyak berbagi waktu, ilmu dan wawasan yang dimiliki kepada penulis dalam penulisan Tugas Akhir ini.



7. Bapak Muhammad Affandes, MT selaku Penguji I Tugas Akhir yang juga telah meluangkan waktunya dan banyak memberikan masukan, ilmu serta wawasan kepada penulis dalam penyusunan Tugas Akhir ini.
8. Bapak Suwanto Sanjaya, ST, M.Kom, selaku Penguji II Tugas Akhir yang telah memberikan masukan dan arahan kepada penulis dalam penyusunan tugas akhir ini.
9. Bapak dan Ibu dosen Jurusan Teknik Informatika yang telah banyak memberikan ilmunya kepada penulis dari awal penulis melaksanakan perkuliahan hingga saat ini.
10. Bapak Deni Karsana, S.S, M.A selaku Validator dan juga sebagai Pakar yang telah membantu penulis dalam melakukan validasi terhadap penelitian Tugas Akhir ini.
11. Terima kasih kepada teman-teman begadang saya yang telah banyak meluangkan waktu dalam membimbing dan memberikan motivasi kepada penulis sehingga dapat menyelesaikan Tugas Akhir ini.
12. Terima kasih kepada Eka Sartika, S.Pd.i yang terus menyemangati serta mendukung penulis selama menyelesaikan penelitian Tugas Akhir sehingga bisa selesai seperti saat ini.
13. Terima kasih kepada teman-teman seperjuangan TIF F 2013 telah menjadi teman penulis semenjak berkuliah, dan juga teman teman Jurusan Teknik Informatika angkatan 2013 yang telah banyak membantu dalam menyelesaikan Tugas Akhir ini.

Semoga laporan ini dapat bermanfaat bagi penulis dan juga pembaca pada umumnya. Masukan, kritik, ataupun saran dapat disampaikan langsung kepada penulis atau dengan alamat email : [surya.atmaja@students.uin-suska.ac.id](mailto:surya.atmaja@students.uin-suska.ac.id)

Akhir kata penulis mengucapkan terimakasih dan selamat membaca.

*Wassalamu'alaikum warahmatullahi wabarakatuh.*

Pekanbaru, Jan 2020

Penulis

## DAFTAR ISI

	Halaman
LEMBAR PERSETUJUAN .....	ii
LEMBAR PENGESAHAN.....	iii
LEMBAR HAK KEKAYAAN INTELEKTUAL.....	iv
LEMBAR PERNYATAAN .....	v
LEMBAR PERSEMBAHAN .....	vi
ABSTRAK.....	vii
ABSTRACT .....	viii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xi
DAFTAR GAMBAR.....	xiv
DAFTAR TABEL.....	xv
DAFTAR SIMBOL .....	xvi
DAFTAR RUMUS.....	xvii
DAFTAR LAMPIRAN .....	xviii
BAB I PENDAHULUAN .....	I-1
1.1 Latar Belakang .....	I-1
1.2 Rumusan Masalah.....	I-4
1.3 Batasan Masalah.....	I-4
1.4 Tujuan Penelitian .....	I-4
1.5 Sistematika Penulisan .....	I-5
BAB II LANDASAN TEORI .....	II-1
2.1 Algoritma .....	II-1
2.2 Stemming .....	II-1
2.3 Bahasa Balantak .....	II-3
2.4 Morfologi Bahasa Balantak.....	II-3
2.4.1 Afiksasi .....	II-3
2.4.1.1 Pembentukan Nomina .....	II-4
2.4.1.2 Pembentukan Verba .....	II-5
2.4.1.3 Pembentukan Adjektiva .....	II-7



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:  
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.  
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

2.4.2 Reduplikasi .....	II-7
2.4.2.1 Reduplikasi Nomina.....	II-7
2.4.2.2 Reduplikasi Verba.....	II-7
2.4.2.3 Reduplikasi Adjektiva .....	II-8
2.5 Perhitungan Akurasi.....	II-8
2.6 Penelitian Terkait .....	II-8
<b>BAB III METODOLOGI PENELITIAN.....</b>	<b>III-1</b>
3.1 Tahapan Penelitian.....	III-1
3.2 Pengumpulan Data .....	III-2
3.3 Analisa .....	III-3
3.3.1 Analisa Kebutuhan Data .....	III-3
3.3.2 Analisa Aturan Tata Bahasa.....	III-3
3.4 Perancangan.....	III-3
3.4.1 <i>Flowchart</i> .....	III-4
3.4.2 <i>Pseudocode</i> .....	III-4
3.4.3 Perancangan Basis Data .....	III-4
3.4.4 Perancangan Antar Muka ( <i>Interface</i> ) .....	III-4
3.5 Implementasi dan Pengujian.....	III-4
3.5.1 Implementasi.....	III-4
3.5.2 Pengujian.....	III-5
3.6 Kesimpulan dan Saran.....	III-5
<b>BAB IV ANALISA DAN PERANCANGAN .....</b>	<b>IV-1</b>
4.1 Analisa .....	IV-1
4.1.1 Analisa Kebutuhan Data .....	IV-1
4.1.2 Analisa Aturan Tata Bahasa Balantak .....	IV-3
4.2 Perancangan.....	IV-7
4.2.1 <i>Flowchart</i> .....	IV-7
4.2.2 <i>Pseudocode</i> .....	IV-13
4.2.3 Perancangan Basis Data .....	IV-14
4.2.4 Perancangan Antar Muka ( <i>Interface</i> ) .....	IV-14
<b>BAB V IMPLEMENTASI DAN PENGUJIAN .....</b>	<b>V-1</b>



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

<b>5.1 Implementasi.....</b>	<b>V-1</b>
<b>5.1.1 Batasan Implementasi.....</b>	<b>V-1</b>
<b>5.1.2 Lingkungan Implementasi.....</b>	<b>V-1</b>
<b>5.2 Implementasi Antar Muka .....</b>	<b>V-2</b>
<b>5.2.1 Tampilan Halaman Utama.....</b>	<b>V-2</b>
<b>5.2.2 Tampilan Halaman <i>Input</i> Kata Berimbuhan .....</b>	<b>V-2</b>
<b>5.2.3 Tampilan Halaman <i>Input</i> Kata.....</b>	<b>V-3</b>
<b>5.2.4 Tampilan Halaman <i>Output</i> Kata .....</b>	<b>V-3</b>
<b>5.2.5 Tampilan Halaman Kata Dasar.....</b>	<b>V-4</b>
<b>5.2.6 Tampilan Halaman Kata Uji.....</b>	<b>V-4</b>
<b>5.3 Pengujian Algoritma <i>Stemming</i> Bahasa Balantak .....</b>	<b>V-5</b>
<b>5.3.1 Pengujian <i>White Box</i> .....</b>	<b>V-5</b>
<b>5.3.2 Pengujian Akurasi.....</b>	<b>V-10</b>
<b>5.4 Analisa Hasil Pengujian.....</b>	<b>V-12</b>
<b>BAB VI PENUTUP .....</b>	<b>VI-1</b>
<b>6.1 Kesimpulan .....</b>	<b>VI-1</b>
<b>6.2 Saran.....</b>	<b>VI-1</b>
<b>DAFTAR PUSTAKA.....</b>	<b>xx</b>

## DAFTAR GAMBAR

Gambar	Halaman
3.1 Tahapan Penelitian.....	III-1
4.1 <i>Flowchart</i> tahapan pemilihan kombinasi terbaik.....	IV-8
4.2 <i>Flowchart</i> Kombinasi A penghapusan imbuhan .....	IV-9
4.3 <i>Flowchart</i> Kombinasi B penghapusan imbuhan .....	IV-10
4.4 <i>Flowchart</i> Kombinasi C penghapusan imbuhan .....	IV-11
4.5 <i>Flowchart</i> Kombinasi D penghapusan imbuhan .....	IV-12
4.6 <i>Pseudocode</i> Algoritma <i>Stemming</i> Bahasa Balantak .....	IV-13
4.7 Rancangan Halaman Utama .....	IV-15
4.8 Rancangan Halaman <i>Input</i> Kata Berimbuhan.....	IV-15
4.9 Rancangan Halaman Kata Dasar .....	IV-16
4.10 Rancangan Halaman Kata Uji .....	IV-17
5.1 Tampilan Halaman Utama .....	V-2
5.2 Tampilan Halaman <i>Input</i> Kata Berimbuhan .....	V-3
5.3 Tampilan Halaman <i>Input</i> kata.....	V-3
5.4 Tampilan Halaman <i>Output</i> Kata.....	V-4
5.5 Tampilan Halaman Kata Dasar .....	V-4
5.6 Tampilan Halaman Kata Uji.....	V-5

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber.

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.  
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

UIN SUSKA RIAU





## DAFTAR TABEL

Tabel	Halaman
2.1 Imbuhan Bahasa Bahasa Balantak.....	II-4
2.2 Penelitian Terkait.....	II-9
4.1 Pembuatan Kata Dasar Pada <i>Microsoft Excel</i> .....	IV-1
4.2 Total Kata Dasar .....	IV-2
4.3 Kata Uji.....	IV-3
4.4 Aturan Tata Bahasa Balantak.....	IV-5
4.5 Variasi Kata Perulangan Bahasa Balantak .....	IV-6
4.6 Rancangan basis data tabel Kamus.....	IV-15
4.7 Rancangan basis data tabel Kamus.....	IV-15
5.1 Proses <i>White Box</i> .....	V-6
5.2 Hasil Pengujian.....	V-10
5.3 Kata Uji Yang Gagal Di <i>Stemming</i> Kombinasi A dan C .....	V-13
5.4 Kata Uji Yang Gagal Di <i>Stemming</i> Kombinasi B dan D .....	V-14

© Hak cipta milik UIN Suska Riau


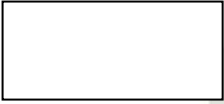
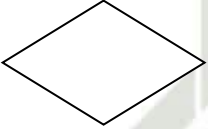


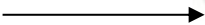
State Islamic University of Sultan Syarif Kasim Riau

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## DAFTAR SIMBOL

### © Flowchart Diagram

Gambar	Keterangan
	Terminator: Simbol terminator (mulai/selesai) merupakan tanda bahwa sistem akan dijalankan atau berakhir.
	Proses: Simbol yang digunakan untuk melakukan pemrosesan data baik oleh user maupun komputer (sistem).
	Verifikasi: Simbol yang digunakan untuk memutuskan apakah valid atau tidak validnya suatu kejadian.
	Data: Simbol yang digunakan untuk mendeskripsikan data <i>input</i> / <i>output</i> yang digunakan.
	Laporan: Simbol yang digunakan untuk menggambarkan laporan
	Arus Data: Simbol yang digunakan untuk menggambarkan arus data di dalam sistem.

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

# DAFTAR RUMUS

Rumus	Halaman
2.1 Perhitungan Akurasi.....	II-20



## Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



## DAFTAR LAMPIRAN

Lampiran	Halaman
A Surat Pernyataan Penunjukan Validator .....	A-1
B <i>Pseudocode</i> Algoritma <i>Stemming</i> Bahasa Balantak .....	B-1
C Data Uji Algoritma <i>Stemming</i> Bahasa Balantak .....	C-1
D Data Penelitian.....	D-1



## BAB I

### PENDAHULUAN

#### 1.1 Latar Belakang

Indonesia merupakan salah satu negara yang mempunyai kepulauan terluas didunia. Tidak hanya luas wilayah, tetapi juga memiliki kekayaan akan seni, budaya dan juga bahasa. Menurut Badan Bahasa Kementerian Pendidikan Nasional mencatat ada 617 bahasa daerah terbesar yang tersebar di beberapa wilayah di Indonesia, 139 bahasa daerah yang terancam punah dan 15 bahasa daerah telah dinyatakan punah (Maulidi, 2016).

Propinsi terbesar di pulau Sulawesi adalah Sulawesi Tengah dengan luas daratan 68,033 kilo meter persegi dan wilayah laut 189,480 kilometer persegi, terletak dibagian barat kepulauan maluku dan bagian selatan Negara Filipina. Secara administratif terbagi dalam Sembilan kabupaten dan satu kota yakni kabupaten Donggala, Parigi moutong, Poso, Morowali, Tojo unauna, Banggai, Banggai kepulauan, Tolitol, Buol dan Sigi serta kota Palu. Jumlah penduduk Sulawesi Tengah 2.875.000 jiwa sesuai sensus penduduk tahun 2007. Bahasa Banggai, bahasa Saluan, dan bahasa Balantak merupakan beberapa bahasa daerah yang terdapat di wilayah Sulawesi Tengah yang masing-masing mempunyai asal lahirnya sendiri. Bahasa Banggai dari daerah Banggai Kepulauan (Kabupaten Banggai Kepulauan), bahasa Saluan dari daerah Saluan dan bahasa Balantak dari daerah Balantak yang keduanya berada di satu kabupaten yaitu Kabupaten Banggai darat (Arfan, 2011).

Bahasa adalah alat komunikasi antar anggota masyarakat yang dihasilkan oleh alat bicara manusia. Setiap manusia dari manapun dia berasal tentu mempunyai bahasa. Jika tidak memiliki bahasa, maka kita tidak lagi dapat menjadi *homo sapiens* atau makhluk yang berpengetahuan. (Ritonga,1:2012). Sedangkan menurut (Walija, 1996), mengungkapkan pengertian bahasa adalah komunikasi yang lengkap dalam menyampaikan sebuah pesan, ide, perasaan, maksud dan juga pendapat untuk orang lain.

Balantak terletak di wilayah paling timur provinsi Sulawesi Tengah, sehingga menjadi salah satu wilayah Sulawesi yang berbatasan langsung dengan

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

laut Maluku. Pada saat ini secara administratif kewilayahan, kecamatan Balantak berada dibawah Kabupaten Banggai dan memiliki batas wilayah sebagai berikut: sebelah selatan berbatasan dengan kecamatan Lamala, bagian utara berbatasan dengan kecamatan Bualemo dan sebelah barat berbatasan dengan kecamatan Masama, sementara bagian timur adalah laut.(Wahyudin, 2017). Sebagai salah satu bahasa daerah yang umum digunakan di Sulawesi Tengah, perlu dibangun sebuah algoritma *stemming text* bahasa Balantak untuk memudahkan para pembaca menterjemahkan teks berbahasa Balantak, baik berupa cerita rakyat, dongeng, dan teks lainnya yang mana algoritma ini bekerja mencari kata dasar dari kata berimbuhan agar kata dasar nya bisa lebih mudah ditemukan di dalam kamus.

*Stemming* adalah proses ekstraksi kata dalam suatu dokumen digital yang bertujuan untuk mendapatkan kata dasarnya dengan mencari prefix, sufiks dan menghapusnya berdasarkan aturan sebuah bahasa. (Winarti, 2017). *Stemming* merupakan sebuah proses pemetaan dan penguraian berbagai bentuk (*variants*) dari suatu kata menjadi bentuk kata dasarnya. Proses *stemming* untuk setiap Bahasa berbeda dengan Bahasa yang lain misal, proses *stemming* Bahasa Inggris dengan Bahasa Indonesia tentunya berbeda karena perbedaan pembentukan dan perubahan kata menjadi bentuk kata lain. Dalam dokumen Bahasa Indonesia proses *stemming* sangat diperlukan sebelum proses *text mining* karena Bahasa Indonesia memiliki *prefixes*, *suffixes*, *infexes* dan *confixes* yang membuat suatu kata dasar dapat berubah menjadi banyak bentuk dan akibatnya membuat pencarian kata dasar menjadi sulit. Begitu juga dengan Bahasa Bali yang memiliki *prefixes*, *suffixes*, *infexes* dan *confixes* seperti Bahasa Indonesia (Nata & Yudiastra, *Stemming Teks Sor-Singgih Bahasa Bali*, 2017).

Algoritma *Stemming* adalah prosedur komputasi untuk mencari asal kata dari suatu kata dalam kalimat dengan cara memisahkan masing-masing kata dari kata dasar dan imbuhan. Algoritma *stemming* untuk bahasa Indonesia yang dikembangkan diantaranya yaitu: Algoritma Nazief dan Andriani, Algoritma *Porter*, serta Algoritma Arifin dan Setiono. (Novitasari, 2016).

Algoritma *Stemming* bahasa daerah sebelumnya sudah pernah diteliti oleh beberapa peneliti antara lain: Penelitian tentang *Implementasi Enhanced Confix*



1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

### *Stripping Stemmer* Untuk Pemisahan Kata Dasar Dalam Dokumen Teks Bahasa

© Batak Angkola-Mandailing yang dilakukan oleh Lisa Noprianti Siregar pada tahun 2018, Penelitian tentang Implementasi Algoritma Porter *Stemmer* untuk Mendapatkan Kata Dasar dalam Dokumen Bahasa Batak Toba dengan Metode Dua Tingkat Morfologi yang dilakukan oleh Tuti Simanjuntak pada tahun 2018, Penelitian tentang Model *Stemming* Berbasis Kamus Untuk Dokumen Berbahasa Sunda yang dilakukan oleh Andhy Purwoko pada tahun 2011, Penelitian tentang Penerapan Modifikasi Metode *Enhanced Confix Stripping Stemmer* Pada Teks Berbahasa Madura yang dilakukan oleh Andi Solihin dkk pada tahun 2013, Penelitian tentang *Stemmer* Bahasa Jawa Ngoko dengan Metode *Affix Removal Stemmers (Rule Based Approach)* yang dilakukan oleh Fatkhul Amin dkk pada tahun 2016, Penelitian tentang *Stemming* bahasa jawa untuk mencari akar kata dalam bahasa jawa dengan aturan analisis kontraktif afiksasi verba yang dilakukan oleh Mat'atus Madia pada tahun 2016, Penelitian tentang *Stemmer* untuk bahasa madura dengan modifikasi metode *enhanced confix stripping stemmer* yang dilakukan oleh Rakhmad Maulidi pada tahun 2016, Penelitian tentang Pengembangan Metode *Stemmer* Untuk Bahasa Bali Dengan Pendekatan *Rule-Based* dan *N-Gram Stemming* yang dilakukan oleh Made Agus Putra Subali pada tahun 2019, Penelitian tentang Aplikasi Kamus Digital Bahasa Indonesia-Karo dengan *Output* Aksara Batak Karo Menggunakan *Enhanced Confix Stripping Stemmer (ECS)* yang dilakukan oleh Upik Purnamawati pada tahun 2014, Penelitian tentang *Stemming* Teks *Sor-Singgih* Bahasa Bali yang dilakukan oleh Gusti Ngurah Mega Nata dan Putu Pande Yudiastra pada tahun 2017, Penelitian tentang Kamus Bahasa Lampung Berbasis Android Dengan Pendekatan *Porter Stemmer* yang dilakukan oleh Nisar & Indera pada tahun 2016 dan penelitian tentang *Stemming Text* Berbahasa Jawa Ngoko-Krama yang dilakukan oleh Kathryn Widhiyanti pada tahun 2007.

Bahasa Balantak merupakan salah satu bahasa yang dipergunakan sehari-hari dalam pergaulan di Kecamatan Balantak dan Lamala, Kabupaten Banggai, Sulawesi Tengah. Pemakai bahasa Balantak ini berjumlah 22.335 orang. Kota Kecamatan Balantak sangat lancar hubungan nya dengan daerah lain. Minat untuk

memajukan pendidikan di daerah itu sangat besar sehingga penduduknya banyak yang bersekolah di tempat lain. Dengan demikian, bahasa Balantak sudah menambah kosa katanya, berupa pinjaman kata, baik yang beradaptasi maupun pinjaman murni. (Rozali, Hente, Mustafa, & P., 1992).

Melihat jumlah pemakai yang relatif besar ini dan kurangnya penelitian tentang bahasa Balantak serta belum adanya penelitian tentang *stemming* bahasa Balantak ini, maka bahasa Balantak perlu diteliti dan dideskripsikan. Berdasarkan latar belakang yang telah dikemukakan tentang bahasa Balantak ini penulis ingin membangun sebuah algoritma *stemming* teks yang berfokus pada bahasa Balantak menggunakan aturan tata bahasa yang hasilnya bermanfaat untuk mengembangkan bahasa daerah itu sendiri, dan diharapkan dapat membantu peneliti selanjutnya untuk mengembangkan ke tahapan normalisasi atau translasi (terjemahan).

## 1.2 Rumusan Masalah

Berdasarkan permasalahan yang timbul pada latar belakang diatas, rumusan masalah yang dapat penulis ambil yaitu bagaimana membangun algoritma *stemming* teks untuk bahasa Balantak menggunakan aturan tata bahasa.

## 1.3 Batasan Masalah

Dalam pembuatan laporan tugas akhir ini mempunyai beberapa batasan masalah dalam penelitian ini, yaitu:

1. Kata dasar diambil dari kamus yang bersumber dari Buku Percakapan Sehari-hari dalam Bahasa Balantak (Peding, Busenitz, & Pontoh, 1991)
2. Aturan yang digunakan sesuai dengan Buku Struktur Bahasa Balantak (Rozali, Hente, Mustafa, & P., 1992)
3. Data yang diperoleh berasal dari cerita rakyat dalam buku Struktur Sastra Lisan Balantak. (Efendi, Kambay, & Tiban, Struktur Sastra Lisan Balantak, 2000)

## 1.4 Tujuan Penelitian

Tujuan dari penelitian yang dilakukan adalah:

1. Membuat algoritma *stemming* teks pada bahasa Balantak.
2. Menguji tingkat akurasi algoritma *stemming* ini dengan menggunakan teks bahasa Balantak untuk memperoleh kata dasarnya.

## 1.5 Sistematika Penulisan

Dalam penulisan laporan Tugas Akhir ini dibuat rencana kerangka kerja penulisan agar dapat dipahami dengan mudah. Sistematika penulisan pada proposal Tugas Akhir ini yaitu:

### **BAB I PENDAHULUAN**

Bab ini menjelaskan tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian dan sistematika penulisan pada laporan Tugas Akhir ini.

### **BAB II LANDASAN TEORI**

Bab ini menjelaskan tentang dasar-dasar teori yang digunakan dalam pembuatan laporan Tugas Akhir ini seperti Pengertian Algoritma, Pengertian *Stemming*, Pengertian Bahasa Balantak, Imbuhan dalam bahasa Balantak dan Penelitian Terkait tentang *stemming* pada bahasa daerah di Indonesia.

### **BAB III METODOLOGI PENELITIAN**

Bab ini menjelaskan tentang metodologi yang dipakai dalam pembuatan laporan Tugas Akhir ini seperti Identifikasi Masalah, Studi Literatur, Pengumpulan Data, Analisa dan Perancangan, Implementasi, serta Kesimpulan dan Saran.

### **BAB IV ANALISA DAN PERANCANGAN**

Bab ini membahas analisa langkah kerja dan perancangan aplikasi yang akan dikembangkan sesuai dengan metodologi yang digunakan.

### **BAB V IMPLEMENTASI DAN PENGUJIAN**

Bab ini akan dibahas implementasi dan pengujian hasil penelitian sesuai yang tertera pada rumusan masalah dan pembahasan.

### **BAB VI PENUTUP**

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber.

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Penutup merupakan jawaban atas rumusan masalah dalam penelitian dan juga intisari dari penelitian serta saran untuk penelitian selanjutnya



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## BAB II

### LANDASAN TEORI

#### 2.1 Algoritma

Algoritma adalah sebuah sistem kerja komputer memiliki perangkat keras, perangkat lunak dan juga *user* yang menjalankan sistem kerja komputer tersebut. Kurang nya salah satu dari 3 sistem tersebut, maka komputer tidak dapat bekerja dan beroperasi. *Software* terdiri dari sebuah susunan program dan juga cara penulisan atau pembuatan program tersebut. Untuk menyusun sebuah program atau aplikasi diperlukan beberapa langkah yang dapat menyelesaikan masalah dan juga tujuan dalam proses pembuatan sebuah *software* atau aplikasi tersebut. Maka algoritma berperan penting dalam melakukan suatu penyusunan program ataupun *syntax* tersebut. (Maulana, 2017)

Algoritma merupakan sebuah susunan yang logis dan juga sistematis untuk menyelesaikan suatu masalah ataupun untuk mencapai suatu tujuan tertentu. Dalam dunia komputer, algoritma berperan penting dalam perancangan dan pembuatan sebuah aplikasi atau *software*. Dalam kehidupan sehari-hari, tanpa kita sadari algoritma sudah masuk dalam kehidupan kita. (Maulana, 2017)

#### 2.2 Stemming

*Stemming* merupakan proses menemukan akar kata (*root*) atau kata dasar dengan memisahkan semua *affix* atau imbuhan yang melekat pada kata tersebut (Indrajaya, 2003). Proses *stemming* digunakan untuk mengubah *term* yang masih melekat dalam *term* tersebut awalan, sisipan, dan akhiran. *Term* tersebut diproses sampai menjadi *term* kata dasar yang mengacu pada bahasa Indonesia yang benar. Proses *stemming* dilakukan dengan cara menghilangkan semua imbuhan dan digunakan untuk mengganti bentuk dari suatu kata menjadi kata dasar dari kata tersebut yang sesuai dengan struktur morfologi bahasa Indonesia yang benar. (Tala, 2003)

*Stemmer* diharapkan mampu mengurangi dimensi data sehingga nantinya akan meningkatkan performansi dari proses kategorisasi. Semakin sedikit dimensi data, maka akan semakin sedikit *rule-rule* untuk menentukan dokumen

terkategorisasi ke suatu kelas sehingga hasil kategorisasi bisa meningkat. © Pengurangan dimensi data pada *stemming* terjadi karena kata-kata yang memiliki kata dasar yang sama dikelompokkan kedalam satu atribut sehingga dimensi data tersebut bisa direduksi. (Amin, Purwatiningtyas, Utomo, Ramadhanu, & Cahya, 2016)

Algoritma *stemming* merupakan salah satu algoritma yang digunakan untuk meningkatkan performa IR caranya yaitu mentransformasikan semua kata dalam teks dokumen ke dalam “*rootword*” atau kata dasarnya. Algoritma *stemming* untuk satu bahasa dan bahasa lainnya berbeda, contohnya bahasa Inggris dengan bahasa Indonesia akan berbeda Algoritma nya karena memiliki morfologi yang berbeda. (Pramudita, 2001)

Pada saat ini Algoritma *stemming* untuk bahasa Indonesia telah dikembangkan oleh beberapa peneliti (Novitasari, 2016) diantaranya Algoritma Nazief dan Andriani, Algoritma *Porter*, serta Algoritma Arifin dan Setiono:

1. Algoritma *stemming* Nazief dan Andriani, Algoritma ini dibuat oleh Bobby Nazief dan Mirna Andriani dari fakultas Ilmu Komputer Universitas Indonesia tahun 1996, seperti diketahui algoritma ini mengacu pada aturan morfologi Bahasa Indonesia yang mengelompokkan imbuhan, pengelompokan ini termasuk imbuhan awalan, imbuhan akhiran, imbuhan sisipan, dan konfiks. Algoritma ini menggunakan kamus kata keterangan yang nantinya digunakan untuk mendapatkan kata dasar dari proses *stemming*.
2. *Porter stemmer for* bahasa Indonesia dikembangkan oleh Fadillah Z. Tala pada tahun 2003 yang berdasar dari English *Porter stemmer* yang dikembangkan oleh W.B Frakes tahun 1992. *Porter stemmer* ini dimodifikasi dan disesuaikan agar sesuai dengan bahasa Indonesia.
3. Algoritma Arifin dan Setiono didahului dengan pembacaan tiap kata dari kata yang ada. Tahapan nya sebagai berikut :
  - a. Pemeriksaan semua kemungkinan bentuk kata. Setiap kata diasumsikan memiliki Awalan (*prefiks*) dan 3 Akhiran (*sufiks*)
  - b. Pemotongan dalam Algoritma ini dilakukan secara berurutan sebagai berikut :



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

AW : AW (Awalan)

AK : AK (Akhiran)

KD : KD (Kata Dasar)

- AW I, hasilnya disimpan pada pe1 (*prefiks 1*)
- AW II, hasilnya disimpan pada pe2 (*prefiks 2*)
- AK I, hasilnya disimpan pada su1 (*sufiks 1*)
- AK II, hasilnya disimpan pada su2 (*sufiks 2*)
- AK III, hasilnya disimpan pada su3 (*sufiks 3*)

c. Akan tetapi apabila sampai pada tahap AK III, belum ditemukan dalam kamus, maka akan dilakukan proses kombinasi.

## 2.3 Bahasa Balantak

Bahasa Balantak adalah bahasa yang digunakan oleh masyarakat yang bermukim di Kecamatan Balantak dan Kecamatan Lamala, Kabupaten Luwuk Banggai, Provinsi Sulawesi Tengah. Nama Balantak selain sebagai bahasa juga digunakan sebagai nama etnis, yaitu etnis Balantak. Masyarakat Balantak memiliki budaya yang tetap terpelihara sampai sekarang sebagai budaya daerah yang mereka banggakan.

Penelitian tentang bahasa Balantak sudah pernah dilakukan oleh Machmud (1986) dengan judul *Abad Banggai Sepintas Kilas* yang lebih membahas pada sejarah dan sastra Balantak. Selain itu juga ada penelitian oleh Rozali et al. (1992) dengan judul *Struktur Bahasa Balantak* yang membahas pada latar belakang sosial budaya masyarakat Balantak, kedudukan, fungsi, nilai, manfaat sastra lisan Balantak, analisis struktur, dan lingkungan penceritaan. (Efendi, Kambay, & Tiban, *Struktur Sastra Lisan Balantak*, 2000)

## 2.4 Morfologi Bahasa Balantak

Proses Morfologi terjadi dalam pembentukan nomina (n), verba (v), dan adjektiva (a), berikut penjelasan nya:

### 2.4.1 Afiksasi

Peristiwa pengimbuhan dalam bahasa Balantak meliputi awalan dan akhiran yang didalam proses nya menggunakan pembentukan nomina (n), verba (v), dan

adjektiva(a). Distribusi imbuhan yang terdapat dalam bahasa Balantak sebagai

©Berikut(Rozali, Hente, Mustafa,& P., 1992) :

**Tabel 2.1 Tabel Imbuhan Bahasa Balantak**

No	Awalan ( <i>prefiks</i> )	Akhiran ( <i>Sufiks</i> )	Gabungan Awalan dan Akhiran ( <i>simulfiks</i> )
1	moN-	-kon	moN-...-kon
2	maN	-i	moN-...-i
3	meN	-an	ni-...-i
4	poN	-on	poN-...-an
5	paN		poN-...-kon
6	peN		poN-...-an
7	ko-		paN...-an
8	ni-		peN-...-an
9	sa-		ko-...-an.

#### 2.4.1.1 Pembentukan Nomina

Dalam bahasa Balantak yang berfungsi membentuk nomina ialah poN-, paN-, peN-, -on, -an, poN-...-an, poN-...-on, peN-...-on, ko-...-an, dan sa-.

- a. Afiks poN-+V, paN-+V, dan peN-+V menyatakan pelaku perbuatan atau alat untuk melakukan perbuatan yang disebut oleh kata dasar. Awalan poN- biasa bervariasi sebagai paN- apabila V berawal dengan fonem /s/ atau /r/; poN- menjadi peN-; apabila V berawal dengan fonem /k/ diikuti oleh /e/.

Contoh :

/koyo/ 'bantai' → /pongkoyo/ 'pembantai'  
/loe/ 'gantong' → /pongoloe/ 'penggantong'  
/ruruki/ 'memilih' → /pangaruruki/ 'pemilih'  
/leot/ 'halang' → /pengeleotan/ 'penghalang'  
/keke/ 'gali' → /pengkekean/ 'penggalian'

- b. Bentuk V+-on 'membentuk nomina abstrak'

Contoh :

/sap/ 'iris' → /sapon/ 'irisan'  
/tulun/ 'tolong' → /tulunon/ 'pertolongan'  
/sarui/ 'tuduh' → /saruiion/ 'tuduhan'  
/bontul/ 'dorong' → /bontulon/ 'dorongan'

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- c. Bentuk ka-+V+-an dan paN-+a-an ‘membentuk nomina abstrak’

Contoh :

/hansur/ ‘hancur’ → /kahansuran/ ‘kehancuran’  
 /idek/ ‘jelek’ → /kaidekan/ ‘kejelekan’  
 /bosi/ ‘busuk’ → /kabosian/ ‘kebusukan’  
 /kaamat/ ‘baik’ → /pankaamatan/ ‘perbaikan’  
 /alus/ ‘halus’ → /panalusan/ ‘penghalusan’

- d. Bentuk sa-+n membentuk nomina yang berarti satu, seluruh, atau seisi.

Contoh :

/lipu/ ‘kampung’ → /salipu/ ‘sekampung’  
 /karun/ ‘karung’ → /sakarung/ ‘sekarung’  
 /rombonan/ ‘rombongan’ → /sarombongan/ ‘serombongan’  
 /ember/ ‘ember’ → /saember/ ‘seember’

#### 2.4.1.2 Pembentukan Verba

Dalam bahasa Balantak imbuhan yang membentuk verba ialah moN-, moN-...-kon, maN-...-kon, ni-, ni-...-kon, dan ni-...-i.

- a. Bentuk moN-+v menyatakan melakukan pekerjaan yang tersebut pada kata dasar. Verba ini tidak transitif.

Contoh :

/kot/ ‘ikat’ → /mongkot/ ‘mengikat’  
 /kolon/ ‘potong’ → /mongkolon/ ‘memotong’  
 /guas/ ‘main’ → /mongguas/ ‘bermain’  
 /ulei/ ‘tuntut’ → /mongulei/ ‘menuntut’  
 /keke/ ‘gali’ → /mongkeke/ ‘menggali’  
 /kait/ ‘nyala’ → /mongkait/ ‘menyala’

- b. Bentuk moN(a)-v+-kon berarti melakukan pekerjaan yang tersebut pada kata dasar. Verba ini adalah transitif.

Contoh :

/ruruki/ ‘pilih’ → /mongarurukikon/ ‘memilihkan’  
 /guas/ ‘main’ → /mongguaskon/ ‘memainkan’  
 /kot/ ‘ikat’ → /mongkotkon/ ‘mengikatkan’



**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

/ator/ 'angkat' → /mengatorkon/ 'mengangkatkan'

/gugur/ 'goyang' → /menggugurkon/ 'menggoyangkan'

/rusi/ 'dapat' → /monrusikon/ 'mendapatkan'

- c. Bentuk v+-konmembuat verba imperatif verba perintah.

Contoh :

/sampur/ 'campur' → /sampurkon/ 'campurkan'

/soba/ 'coba' → /sobakon/ 'cobakan'

/rusi/ 'dapat' → /rusikon/ 'dapatkan'

/loe/ 'gantung' → /loekon/ 'gantungkan'

/inum/ 'minum' → /inumkon/ 'minumkan'

- d. Bentuk ni-v membentuk verba pasif

Contoh :

/ator/ 'angkut' → /niator/ 'diangkut'

/gugur/ 'goyang' → /nigugur/ 'digoyang'

/inti/ 'kenal' → /niinti/ 'dikenal'

/koyo/ 'bantai' → /nikoyo/ 'dibantai'

/layak/ 'goda' → /nilayak/ 'digoda'

- e. Bentuk moN-+n membentuk verba aktif yang berarti membuat atau melakukan seperti yang tersebut pada kata dasar.

Contoh :

/gambar/ 'gambar' → /monggambar/ 'menggambar'

/bose/ 'dayung' → /mombose/ 'mendayung'

/sonto/ 'contoh' → /monsonto/ 'mencontoh'

/buani/ 'jala' → /mombuani/ 'menjala'

/paat/ 'pahat' → /monpaat/ 'memahat'

- f. Bentuk moN-+a(i) membentuk verba yang berarti membuat atau membuat lebih seperti yang tersebut pada kata dasar; moN-+a membuat lebih seperti yang tersebut pada kata dasar; moN-+a membuat bentuk verba tidak transitif, sedangkan moN-+a(i) membentuk verba transitif.

Contoh :

/ampudu/ 'pendek' → /mongampudui/ 'memendekkan'

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

/alalom/ 'dalam' → /mongalalom/ 'memperdalam'

/alus/ 'halus' → /mongalusi/ 'memperhalus'

/hansur/ 'hancur' → /monghansur/ 'menghancur'

/balaki/ 'besar' → /mom-balaki/ 'membesarkan'

### 2.4.1.3 Pembentukan Adjektiva

Dalam bahasa Balantak imbuhan yang membentuk adjektiva ialah -na. Bentuk a-na menyatakan tingkat perbandingan yang lebih (komparatif).

Contoh :

/alayo/ 'panjang' → /alayo?na/ 'lebih panjang'

/itiu?/ 'kecil' → /itiu?na/ 'lebih kecil'

/molono?/ 'gemuk' → /malono?na/ 'lebih gemuk'

/babulak/ 'putih' → /babulakna/ 'lebih putih'

### 2.4.2 Reduplikasi

Berdasarkan kata dasarnya reduplikasi atau perulangan dalam bahasa Balantak dapat dibedakan atas perulangan seluruhnya, perulangan sebagian berkombinasi dengan afiksasi, perulangan dengan peluluhan fonem, dan perulangan yang hanya mempunyai arti apabila diulang. Kata dasar yang diulang dapat berupa nomina, verba, dan adjektiva.

#### 2.4.2.1 Reduplikasi Nomina

Dalam perulangan bahasa Balantak ditemukan bentuk kata dasar yang diulang seluruhnya dan sebagian arti 'mainan' atau 'tidak sesungguhnya', sebagian kata dasar mempunyai arti yang berbeda dari perulangan nya dan sebagian lagi hanya mempunyai arti apabila diulang.

Contoh:

/ajaran/ 'kuda' → /ajaraajaran/ 'kuda-kuda'

/buani/ 'jala' → /buabuani/ 'jala-jalaan'

/pac/ 'beras' → /paepac/ 'gumpalan kecil'

#### 2.4.2.2 Reduplikasi Verba

Dalam bahasa balantak terdapat perulangan seluruh kata dasar dan perulangan sebagian kata dasar verba. Perulangan sebagian dapat berkombinasi

dengan imbuhan. Perulangan verba pada umumnya menunjukkan pekerjaan yang berulang-ulang, menyatakan intensitas atau melemahkan arti.

Contoh:

/sampur/ 'campur' → /sampursampur/ 'campur-campur'

/sankul/ 'cangkul' → /sankulsankul/ 'cangkul sedikit'

/tulun/ 'tolong' → /montuluntulun/ 'bantu-membantu'

### 2.4.2.3 Reduplikasi Adjektiva

Perulangan adjektiva dalam bahasa Balantak dapat berupa perulangan seluruh dan sebagian kata dasar. Pada perulangan sebagian hanya unsur kata kedua yang diulang seluruhnya. Arti perulangan adjektiva itu menyatakan 'intensitas' atau 'sedikit', perulangan sebagian dapat berimbuhan

Contoh:

/kodi/ 'sedikit' → /kodikodi/ 'sebentar sekali'

/aloo/ 'jauh' → /alooaloo/ 'sedikit jauh'

/karani/ 'dekat' → /karakarani/ 'agakdekat'

/rawat/ 'berat' → /marawarawat/ → 'berat-berat'

## 2.5 Perhitungan Akurasi

Perhitungan akurasi dilakukan untuk mengetahui tingkat akurasi dari hasil klasifikasi, dengan cara menghitung jumlah *record* uji yang kelasnya diprediksi secara tepat. Dapat dilihat pada persamaan dibawah ini (Parvin et al, 2008).

$$\text{Total Akurasi} = \frac{\text{Jumlah Prediksi Benar}}{\text{Jumlah Total Prediksi}} \times 100 \quad \dots(2.1)$$

Jumlah prediksi benar adalah jumlah *record* data uji yang diprediksi kelasnya menggunakan metode klasifikasi dan hasilnya sama dengan kelas sebenarnya. Sedangkan jumlah total prediksi adalah jumlah keseluruhan *record* yang diprediksi kelasnya (seluruh data uji).

## 2.5 Penelitian Terkait

Berikut ini beberapa penelitian terkait tentang algoritma *stemming text* pada bahasa daerah:



Tabel 2.2 Penelitian Terkait

No	Peneliti (Tahun)	Judul	Keterangan
1	Fatkhul Amin dkk (2016)	<i>Stemmer Bahasa Jawa Ngoko dengan Metode Affix Removal Stemmers (Rule Based Approach)</i>	<i>Stemmer Bahasa Jawa Ngoko</i> bisa membuat kata dasar jawa ngoko dengan benar sebanyak 62% atau 21 dari 34.
2	Andhy Purwoko (2011)	Model <i>Stemming</i> Berbasis Kamus Untuk Dokumen Berbahasa Sunda	Pada pengujian seluruh kata bahasa sunda yang memiliki 10.416 Kata, <i>stemming</i> yang dirancang menghasilkan jumlah pengurangan kata sebesar 49.31%. Dari jumlah kata hasil <i>stemming</i> (5.136 Kata), Kata yang sesuai harapan berjumlah 4.693 kata dan 443 kata-kata yang tidak sesuai harapan. Dari data ini dapat dihitung bahwa hasil akurasi <i>stemming</i> yang diperoleh sebesar 91.38%
3	Andi Solihin dkk (2013)	Penerapan Modifikasi Metode <i>Enhanced Confix Stripping Stemmer</i> Pada Teks Berbahasa Madura	penelitian ini hanya sebatas menggunakan bahasa Madura dialek Bangkalan dan persentase kata pada bahasa Madura yang berhasil diuji adalah 95,75% dari 400 kata yang telah diuji
4	Mar'atus Mada(2016)	<i>Stemming</i> bahasa jawa untuk mencari akar kata dalam bahasa jawa dengan aturan analisis kontraktif afiksasi verba	Akurasi yang dihasilkan dari pengujian algoritma <i>Stemming</i> bahasa jawa menggunakan aturan analisis kontrasif afiksasi verba sebesar 92% dari jumlah data 16.745 kata bahasa jawa.
5	Rakhmad maulidi (2016)	<i>Stemmer</i> untuk bahasa madura dengan modifikasi metode <i>enhanced confix stripping stemmer</i>	<i>Stemming</i> merupakan sebuah teknik yang penting didalam informasi retrieval, yang mana teknik tersebut dapat digunakan untuk proses klasifikasi, translasi atau ringkasan dokumen secara otomatis.
6	Lisa Noprianti Siregar(2018)	<i>Implementasi Enhanced Confix Stripping Stemmer</i> Untuk Pemisahan Kata Dasar Dalam Dokumen Teks Bahasa Batak Angkola-Mandailing	Penelitian ini menghasilkan hasil uji akurasi sebesar 87%, sistem <i>stemming</i> tidak dapat menyelesaikan semua kata dasar, kecuali kata tersebut diawali dengan huruf T dan P.

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Peneliti (Tahun)	Judul	Keterangan
7	Made Agus Putra Subali (2019)	Pengembangan Metode <i>Stemmer</i> Untuk Bahasa Bali Dengan Pendekatan <i>Rule-Based</i> dan <i>N-Gram Stemming</i>	Berdasarkan pengujian yang dilakukan untuk 10 <i>query</i> dinyatakan benar dari 15 <i>query</i> . Metode yang telah diusulkan memperoleh akurasi <i>stemming</i> lebih baik yaitu sebesar 85,83% dari metode yang terdahulu yang hanya memperoleh akurasi sebesar 56,67%, untuk pengujian 20 dokumen akurasi yang dihasilkan lebih baik yaitu sebesar 79,26% dari metode yang terdahulu yang hanya memperoleh akurasi sebesar 56,13%. Sedangkan untuk pengujian 5 dari 15 <i>query</i> menggunakan metode <i>N-Gramstemming</i> dapat mengenali kata-kata yang berafiks diluar <i>rules</i> , dimana terdapat kata berafiks yang ditulis salah dan juga adanya kata yang mengalami proses disimilasi.
8	Upik Purnamawati (2014)	Aplikasi Kamus Digital Bahasa Indonesia-Karo dengan <i>Output</i> Aksara Batak Karo Menggunakan <i>Enhanced Confix Stripping Stemmer (ECS)</i>	Aplikasi Kamus bahasa Indonesia-Karo ini bukan hanya sekedar pencarian kata dalam bahasa Indonesia-Karo tetapi juga sebagai bahan pembelajaran untuk mengenal aksara Batak Karo. Algoritma ECS mampu mencari kata baku secara efektif dari imbuhan khususnya bahasa Indonesia dan selanjutnya dapat diterjemahkan kedalam bentuk bahasa Batak Karo
9	Tuti Simanjuntak(2018)	Implementasi Algoritma Porter <i>Stemmer</i> untuk Mendapatkan Kata Dasar dalam Dokumen Bahasa Batak Toba dengan Metode Dua Tingkat Morfologi	Dari hasil pengujian pada 20 dokumen bahasa Batak Toba, algoritma Porter <i>stemmer</i> dengan dua tingkat morfologi ini dapat menghasilkan tingkat akurasi 87,503%

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Peneliti (Tahun)	Judul	Keterangan
10	Gusti Ngurah Mega Nata dan Putu Pande Yiudiasra (2017)	<i>Stemming Teks Sor-Singgih Bahasa Bali</i>	Pada pengujian ini jumlah kata sor-singgih yang digunakan sejumlah 357 kata pada satu dokumen bahasa Bali. Dari hasil pengujian 85% kata di <i>stemming</i> dengan benar.
11	Nisar & Indera (2016)	Kamus Bahasa Lampung Berbasis Android Dengan Pendekatan <i>Porter Stemmer</i>	Kamus ini memfasilitasi input yang berupa kata dasar dan juga kata yang berimbuhan. <i>Output</i> yang dihasilkan pada aplikasi kamus ini berupa teks dan juga suara yang merupakan hasil dari terjemahan
12	Kathryn Widhiyanti (2007)	<i>Stemming Text</i> Berbahasa Jawa Ngoko-Krama	Dalam penelitian ini, algoritma <i>stemming text</i> yang digunakan hanya berfokus terhadap bahasa jawa ngoko dan bahasa jawa krama secara umum (yang sering digunakan). Dari hasil pengujian algoritma <i>stemming text</i> pada bahasa jawa ngoko dan krama ini sudah menghasilkan kata dasar yang sesuai dengan aturan yang didapat dan juga kata dalam kamus



#### Hak Cipta Dilindungi Undang-Undang

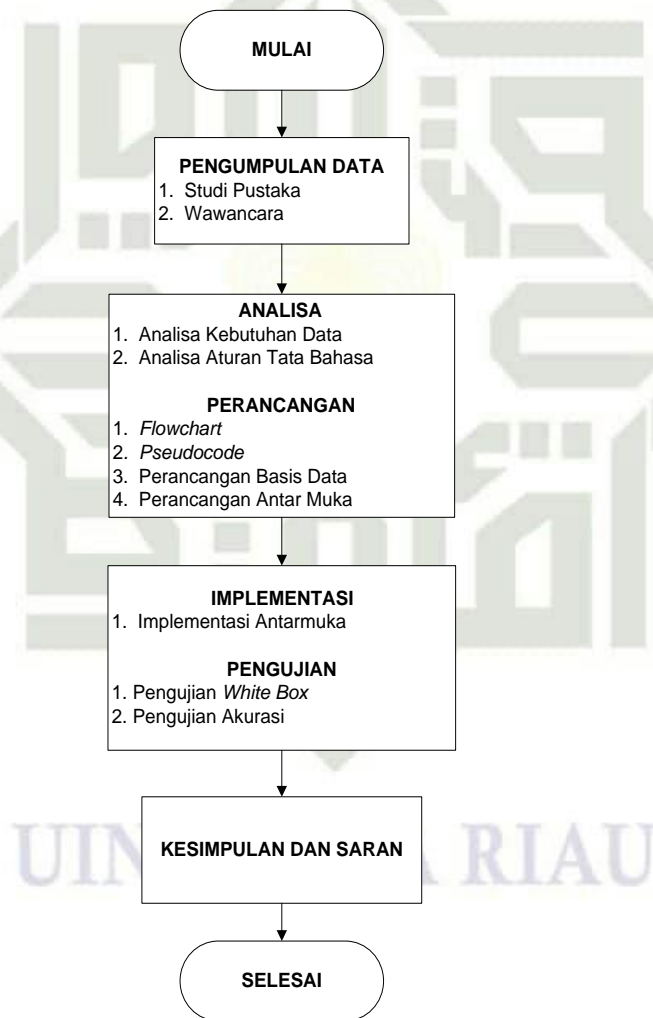
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## BAB III

### METODOLOGI PENELITIAN

#### 3.1 Tahapan Penelitian

Metodologi penelitian bertujuan untuk menguraikan kegiatan yang akan dilaksanakan selama kegiatan penelitian berlangsung. Dalam metodologi penelitian terdapat beberapa tahapan yang akan dilakukan untuk menyelesaikan penelitian ini. Tahapan penelitian bisa dilihat pada gambar berikut:



Gambar 3.1 Tahapan Penelitian

Gambar 3.1. di atas merupakan metodologi yang akan dilakukan oleh penulis. Metodologi penelitian bertujuan untuk menguraikan kegiatan yang akan dilaksanakan selama kegiatan penelitian berlangsung. Dari Gambar 3.1. terdapat beberapa tahapan yang akan dilakukan untuk menyelesaikan kasus pada penelitian tugas akhir ini yang akan meliputi pengumpulan data, analisa, perancangan, implementasi dan pengujian, serta kesimpulan dan saran.

### 3.2 Pengumpulan Data

Tahapan pengumpulan data merupakan tahapan yang paling penting dalam penelitian ini, dimana proses ini dilakukan untuk memperoleh informasi-informasi atau data-data terkait kasus yang menjadi permasalahan dalam laporan tugas akhir ini. Data-data yang diperlukan dalam penelitian ini berasal dari:

#### 1. Studi Pustaka

Melakukan penelitian tahap awal untuk mencari informasi-informasi awal mengenai penelitian-penelitian yang pernah dilaksanakan sebelumnya, yang berhubungan dengan algoritma *stemming*. Studi pustaka dilakukan untuk memperoleh teori-teori yang mendasar mengenai materi yang berhubungan dalam penelitian kali ini yaitu dengan cara mempelajari buku, artikel, jurnal dan kebutuhan data yang diperoleh berasal dari Buku Percakapan Sehari-hari dalam Bahasa Balantak, Morfologi bahasa Balantak berasal dari Buku Struktur Bahasa Balantak dan data uji yang diperoleh berasal dari buku Struktur Sastra Lisan Balantak.

#### 2. Wawancara

Wawancara atau tanya jawab dengan pakar bahasa Balantak sekaligus permohonan penunjukan validator dalam melakukan penelitian Tugas Akhir ini. Dalam melakukan validasi, validator melakukan proses pemisahan kata berimbuhan dari kata uji yang telah dibuat dan diberikan ke validator sebelumnya untuk mendapatkan kata dasar sekaligus memberikan arti dari kata berimbuhan tersebut. (Biodata validator dapat dilihat pada lampiran A).

### 3.3 Analisa

Setelah mendapatkan informasi melalui studi pustaka, wawancara dan mengumpulkan data penelitian, selanjutnya dilakukan tahapan analisa sehingga akan dapat diketahui tentang gambaran yang jelas mengenai penelitian yang dilakukan. Analisa dilakukan terhadap data-data yang telah diperoleh dan diproses menggunakan metode aturan tata bahasa sesuai dengan morfologi dan sintaksis bahasa Balantak. Adapun rincian analisisnya sebagai berikut:

#### 3.3.1. Analisa Kebutuhan Data

Tahapan ini merupakan tahapan menganalisa data penelitian yang telah dikumpulkan dan akan diolah menggunakan metode aturan tata bahasa, sekaligus pembuatan kata dasar dan kata uji. Adapun data yang digunakan bersumber dari:

1. Kata dasar yang diambil bersumber dari Buku Percakapan Sehari-hari dalam Bahasa Balantak (Peding, Busenitz, & Pontoh, 1991)
2. Aturan yang digunakan sesuai dengan Buku Struktur Bahasa Balantak (Rozali, Hente, Mustafa, & P., 1992)
3. Data yang diperoleh berasal dari cerita rakyat dalam buku Struktur Sastra Lisan Balantak. (Efendi, Kambay, & Tiban, Struktur Sastra Lisan Balantak, 2000).

Data yang diperoleh dari buku Percakapan Sehari-hari dalam Bahasa Balantak dan buku Struktur Sastra Lisan Balantak yang melampirkan percakapan sehari-hari bahasa dan cerita rakyat bahasa Balantak ini berupa teks yang telah diketik ulang menjadi data elektronik sehingga data tersebut dapat digunakan dalam tahap pengujian dan data disimpan ke dalam *database*.

#### 3.3.2. Analisa Aturan Tata Bahasa Balantak

Pada tahapan ini langkah yang dilakukan adalah menganalisa aturan tata bahasa Balantak. Algoritma *stemming* bahasa Balantak ini akan dianalisa berdasarkan aturan dari tata bahasa Balantak. Analisa yang dilakukan yaitu:

1. Mengumpulkan aturan-aturan pembentukan kata dengan imbuhan berdasarkan tata bahasa Balantak.
2. Merumuskan aturan-aturan yang menghilangkan imbuhan berdasarkan aturan pembentukan kata dengan imbuhan bahasa Balantak.
3. Menentukan urutan aturan-aturan menghilangkan imbuhan.



1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

### 3.4 Perancangan

Pada tahapan ini langkah yang dilakukan adalah membuat *Flowchart*, *pseudocode* dan juga basis data yang dibutuhkan dalam membangun algoritma *stemming* teks bahasa Balantak berdasarkan hasil dari analisa dan juga data yang telah didapat pada tahapan sebelumnya.

#### 3.4.1. Flowchart

Pada tahapan ini perancangan dengan menggunakan *flowchart* digunakan untuk menggambarkan bagaimana alur dari sebuah proses algoritma yang dibuat.

#### 3.4.2. Pseudocode

Pada tahapan ini berisi langkah-langkah algoritma *stemming* pada bahasa Balantak. Tahapan langkah-langkah algoritma *stemming* ditulis dalam bentuk *pseudocode*.

#### 3.4.3. Perancangan Basis Data

Pada tahapan ini berisi perancangan basis data yang merupakan tahapan pembuatan tabel yang dibutuhkan untuk membuat algoritma *stemming* pada bahasa Balantak.

#### 3.4.4. Perancangan Antar Muka (*Interface*)

Pada tahapan ini berisi perancangan antar muka yang merupakan tahapan pembuatan rancangan tampilan algoritma *stemming* pada bahasa Balantak.

### 3.5 Implementasi dan Pengujian

Pada tahapan ini terdapat implementasi yang merupakan penerapan dari hasil perancangan dan juga pembuatan algoritma yang dilakukan sebelumnya dan pengujian dilakukan untuk mengetahui tingkat akurasi algoritma *stemming* teks bahasa Balantak.

#### 3.5.1. Implementasi

Pada tahapan implementasi ini komponen *hardware* dan *software* yang digunakan adalah:

*Hardware* yang dibutuhkan:

1. *Processor* : Intel Core i3 CPU @1.80 GHz
2. *Memory* : 2048 MB

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3. *Hard Disk* : 500 GB

*Software* yang dibutuhkan:

1. Sistem Operasi : Microsoft Windows 7 32-bit
2. *Web Server* : Apache
3. *Browser* : Google Chrome
4. Bahasa Pemrograman : PHP
5. *Tool* : *Sublime text*, Microsoft Office Visio 2007
6. *Database* : DBMS MySQL

### 3.5.2. Pengujian

Pada tahapan ini dilakukan pengujian *white box* untuk menjelaskan kinerja logika yang dibuat pada algoritma *stemming* untuk mengecek apakah algoritma *stemming* sudah berjalan sesuai dengan aturannya dan pengujian algoritma dilakukan dengan tujuan apakah algoritma *stemming* teks yang dibuat sudah sesuai berdasarkan hasil analisa yang dibuat dan untuk mengukur tingkat akurasi algoritma dengan hasil yang telah ditentukan sebelumnya dengan cara membandingkan hasil algoritma *stemming* dengan hasil yang ditentukan sebelumnya oleh pakar bahasa Balantak. Pengujian ini dilakukan dengan cara *menginputkan* data uji berupa percakapan sehari-hari dan cerita rakyat bahasa Balantak kedalam sistem kemudian sistem akan memproses dan menampilkan apakah kata tersebut ada didalam *database* dan juga menampilkan hasil akurasi dari *stemming* tersebut.

### 3.6 Kesimpulan dan Saran

Pada tahapan ini merupakan kesimpulan dan juga saran yang dapat diambil dari penelitian ini. Kesimpulan adalah hasil dari penelitian yang telah dilakukan dan juga tingkat akurasi dari algoritma *stemming* teks bahasa Balantak yang didapat dari hasil pengujian algoritma yang telah dibuat. Saran merupakan masukan yang didapat dari pembaca agar dapat melakukan pengembangan lebih lanjut berdasarkan hasil penelitian yang sudah dilakukan.

## BAB VI

### PENUTUP

#### 6.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat diambil kesimpulan sebagai berikut:

1. Algoritma *stemming* pada bahasa Balantak berhasil dibuat sesuai dengan aturan tata bahasa Balantak berdasarkan aturan pembentukan imbuhan yang diperoleh pada morfologi dan sintaksis bahasa Balantak.
2. Algoritma *stemming* yang telah dibuat memperoleh tingkat akurasi tertinggi menggunakan Kombinasi B sebesar 98,39% dengan jumlah 310 kata uji.
3. Kata uji yang salah didapat karena mempunyai 2 kata dasar dari 1 kata berimbuhan untuk semua kombinasi.

#### 6.2 Saran

Adapun saran untuk penelitian selanjutnya tentang bahasa Balantak ini yaitu:

1. Algoritma *stemming* dapat dilanjutkan ke tahapan *POS tagging*, dengan membedakan kelas kata maka akan menambah akurasi karena bisa membedakan antara kata kerja dan kata benda agar bisa mengetahui kata yang bisa ditambahkan imbuhan maupun yang tidak, hal ini dilakukan agar penelitian ini lebih baik lagi dikarenakan kata uji yang salah didapat karena kata yang seharusnya tak memiliki imbuhan tidak dipisahkan.
2. Algoritma *stemming* bahasa Balantak bisa dilanjutkan ke tahapan normalisasi pada bahasa Balantak karena *library* sudah lengkap dan bisa digunakan untuk pengecekan kata yang ingin dinormalisasi.



## DAFTAR PUSTAKA

- Afuan, L. (2013). Stemming dokumen teks bahasa indonesia menggunakan algoritma porter. *Jurnal Telematika Vol. 6* , 34-40.
- Amin, F., Purwatiningsy, Utomo, P., Ramadhanu, S., & Cahya, S. E. (2016). *Stemmer Bahasa Jawa Ngoko dengan Metode Affix Removal Stemmers (Rule Based Approach)*. Semarang: Fakultas Teknologi Informasi Universitas Stikubank ( Unisbank) Semarang.
- Astuti, I. P. (2011). *Aplikasi Tashrif Istilahi pada Ilmu Sharaf dengan Menggunakan Metode Stemmer Berbasis Web*. Jakarta: Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam NEgeri Syarif Hidayatullah.
- Efendi, Kambay, S. B., & Tiban, A. R. (2000). *Struktur Sastra Lisan Balantak*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Madia, M. (2016). *Stemming bahasa jawa untuk mencari akar kata dalam bahasa jawa dengan aturan analisis kontraktif afiksasi verba*. Malang: Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim.
- Maulidi, R. (2016). Stemmer untuk Bahasa Madura dengan Modifikasi Metode Enhanced Confix Stripping Stemmer. 12-15.
- Nopiyanti, D., & Sekarwati, K. A. (2014). Aplikasi Pencarian Kata Dasar Dokumen Berbahasa Indonesia dengan Metode Stemming Porter Menggunakan PHP & MySQL. *Prosiding Seminar Ilmiah Nasional Komputer dan Sistem Intelijen (KOMMIT 2014) Universitas Gunadarma Vol.8* , 215-222.
- Novitasari, D. (2016). Perbandingan Algoritma Stemming Porter dengan Arifin Setiono untuk menentukan tingkat ketepatan Kata Dasar. *Jurnal String Vol.1 No.2* , 120-129.
- Peding, W., Busenitz, R. L., & Pontoh, Z. (1991). *Percakapan sehari-hari dalam Bahasa Balantak*. Palu, Sulawesi Tengah: The Summer Institute of Linguistics, Inc.
- Pramudita, H. R. (2001). Penerapan Algoritma Stemming Nazief & Adriani dan similarity pada penerimaan judul Thesis. *Jurnal Ilmiah DASI Vol.15 No. 04* ,

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Ritonga, Parlaungan dkk. 2012. *Bahasa Indonesia Praktis*. Medan: Bartong Jaya. 15-19.

Rosidin, M., Fadlil, A., & Yudhana, A. (2018). Perancangan Sistem Penentuan Kelas Kata Berimbuhan Menggunakan Algoritma Porter Stemmer. *Seminar Nasional Informatika UPN "veteran" Yogyakarta* , 108-114.

Rozali, L., Hente, A., Mustafa, D., & P. Nurhayati. (1992). *Struktur Bahasa Balantak*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa Departemen Pendidikan dan Kebudayaan.

Simanjuntak, T. (2018). *Implementasi Algoritma Porter Stemmer untuk Mendapatkan Kata Dasar Dalam Dokumen Bahasa Batak Toba dengan Metode Dua Tingkat Morfologi*. Medan: Program Studi S1 Teknologi Informasi Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara.

Siregar, L. N. (2018). *Implementasi Enhanced Confix Stripping Stemmer Untuk Pemisahan Kata Dasar Dalam Dokumen Teks Bahasa Batak Angkola-Mandailing*. Medan: Program Studi S1 Teknologi Informasi Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara.

Tala, F. Z. (2003). *A Study of Stemming Effects on Information Retrieval in Bahasa Indonesia*. Institut for Logic, Language and Cumputation Universiteit van Amsterdam The Netherlands.

Utomo, M. S. (2013). Implementasi Stemmer Tala pada Aplikasi Berbasis Web. *Jurnal Teknologi Informasi Dinamik Volume 18* , 41-45.

Walija. (1996). *Bahasa Indonesia Dalam Perbincangan*. Jakarta: IKIP.

Wiguna, P. B., & Hantono, B. S. (2013). Peningkatan Algoritma Porter stemmer Bahasa Indonesia berdasarkan Metode Morfologi dengan Mengaplikasikan 2 Tingkat Morfologi dan Aturan Kombinasi Awalan dan Akhiran. *JNTETI, Vol. 2* , 1-6.

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LAMPIRAN A

### SURAT PERNYATAAN PENUNJUKAN VALIDATOR

Transkrip ini merupakan surat pernyataan penunjukan validator dari Balai Adat bahasa Sulawesi Tengah khusus bahasa Balantak yang bertujuan untuk membantu memvalidasi data uji yang diperoleh dalam melakukan penelitian Tugas Akhir ini.







**Hak Cipta Dilindungi Undang-Undang**

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**SURAT PERNYATAAN**

Saya yang bertanda tangan dibawah ini,

Nama : Deni Karsana, S.S, M.A.  
 NIP : 196905142001121001  
 Pangkat : Pembina / (IV/a)  
 Jabatan : Peneliti Ahli Madya / (IV/a)  
 Alamat : Jl. Untad Elumi Roviga Kelurahan Fondo Kec.  
 Mantikulore, Kota Palu Sulawesi Tengah  
 No Hp : 081342656211

Dengan ini menyatakan bahwa saya bersedia menjadi Validator dari Penelitian

Nama : Surya Atmaja  
 NIM : 11351101946  
 Judul Penelitian: Algoritma *Stemming* teks Bahasa Balantak  
 Berbasis Aturan Tata Bahasa  
 Instansi : UIN SUSKA Riau  
 Jurusan : Teknik Informatika

Demikian surat pernyataan ini saya sampaikan untuk digunakan seperlunya. Atas perhatian, saya mengucapkan terimakasih.

Palu, Agustus 2020

(Deni Karsana, S.S, M.A.)

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LAMPIRAN B

### PSEUDOCODE ALGORITMA STEMMING BAHASA BALANTAK

Transkrip ini merupakan *pseudocode* yang digunakan dalam membangun algoritma *stemming* teks pada bahasa Balantak.





# ALGORITMA STEMMING TEKS BALANTAK

```

FUNCTION cek_kamus (_KATA)

    IF _Kata terdapat dikamus THEN

        RETURN TRUE

    ELSE

        RETURN FALSE

    END IF

END FUNCTION

FUNCTION Cek_Awalan(_KATA, _AWALAN)

    If _kata memiliki _AWALAN THEN

        RETURN TRUE

    ELSE

        RETURN FALSE

    END IF

END FUNCTION

FUNCTION Cek_Akhiran(_KATA, _AKHIRAN)

    IF _Kata memiliki _AKHIRAN THEN

        RETURN TRUE

    ELSE

        RETURN FALSE

    END IF

END FUNCTION

FUNCTION Cek_Perulangan(_KATA, _KATA_PERULANGAN)

    IF _Kata memiliki _KATA_PERULANGAN THEN

        RETURN TRUE

    ELSE

        RETURN FALSE

    END IF

END FUNCTION
  
```





```

FUNCTION Stemming(_KATA)

    _KATA_SETELAH_DIHAPUS = _KATA

    _DAFTAR_AWALAN = ('maN-', 'meN', 'moN', 'paN', 'peN', 'poN', 'ko', 'ni', 'sa')

    _DAFTAR_AKHIRAN = ('-kon', '-an', 'i', 'on')

    _KATA_PERULANGAN = ('-')

    IF !Cek_Kamus(_KATA) THEN

        IF Cek_Awalan(_KATA, _DAFTAR_AWALAN) THEN

            _KATA_SETELAH_DIHAPUS = Hapus_Awalan(_KATA, _DAFTAR_AWALAN)

            IF !Cek_Kamus(_KATA_SETELAH_DIHAPUS) THEN

                IF Cek_Akhiran(_KATA_SETELAH_DIHAPUS, _DAFTAR_AKHIRAN) THEN

                    _KATA_SETELAH_DIHAPUS = Hapus_Akhiran(_KATA, _DAFTAR_AKHIRAN)

                    IF !Cek_Kamus(_KATA_SETELAH_DIHAPUS) THEN

                        IF Cek_Perulangan(_KATA_SETELAH_DIHAPUS,

                            KATA_PERULANGAN), THEN

                            _KATA_SETELAH_DIHAPUS = Hapus_Perulangan(_KATA,

                                _KATA_PERULANGAN)

                            END IF

                        END IF

                    END IF

                END IF

            END IF

        END IF

        RETURN _KATA_SETELAH_DIHAPUS

    ELSE

        RETURN _KATA

    END IF

END FUNCTION

```

## LAMPIRAN C

### DATA UJI ALGORITMA *STEMMING* BAHASA BALANTAK

Transkrip ini merupakan data uji yang diperoleh serta hasil dari algoritma *stemming* teks bahasa Balantak dengan membandingkan hasil dari validator bahasa Balantak.

No	Kata Berimbuhan	Hasil Validator	Hasil Algoritma	Hasil
1	alai	ala	ala	Benar
2	alakon	ala	ala	Benar
3	alaon	ala	ala	Benar
4	mangala	ala	gala	Salah
5	pangala	ala	gala	Salah
6	pangalaan	ala	gala	Salah
7	manganam	anam	anam	Benar
8	angeon	ange	ange	Benar
9	mangange	ange	ange	Benar
10	mangasaqi	asaq	gasaq	Salah
11	asokon	asok	asok	Benar
12	mangasok	asok	asok	Benar
13	niasok	asok	asok	Benar
14	pangasok	asok	asok	Benar
15	pangasokan	asok	asok	Benar
16	mangator	ator	ator	Benar
17	mambala	bala	bala	Benar
18	mambalukkon	baluk	baluk	Benar
19	kobambalan	bambal	bambal	Benar
20	manbangkali	bangkali	bangkal	Salah
21	mabarang	barang	barang	Benar
22	belaon	bela	bela	Benar
23	bobokon	bobok	bobok	Benar
24	bobunogan	bobunoq	bobunoq	Benar
25	boitan	boit	boit	Benar
26	momboiti	boit	boit	Benar
27	koboliiian	bolii	bolii	Benar
28	mobolos	bolos	bolos	Benar
29	mombolos	bolos	bolos	Benar
30	nibubut	bubut	bubut	Benar

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Berimbuhan	Hasil Validator	Hasil Algoritma	Hasil
31	pandaat	daat	daat	Benar
32	mandagai	daga	daga	Benar
33	dodongoan	dodongo	dodongo	Benar
34	mondoloq	doloq	doloq	Benar
35	garedaan	gareda	gareda	Benar
36	ilion	ili	ili	Benar
37	mongimputi	imput	imput	Benar
38	inumon	inum	inum	Benar
39	kaanon	kaan	kaan	Benar
40	mangkaan	kaan	kaan	Benar
41	kamburon	kambur	kambur	Benar
42	pakanaqkon	kanaq	kanaq	Benar
43	mangkarei	karei	karei	Benar
44	karison	karis	karis	Benar
45	mangkaris	karis	karis	Benar
46	kauton	kaut	kaut	Benar
47	mangkaut	kaut	kaut	Benar
48	mokaut	kaut	kaut	Benar
49	kekeon	keke	keke	Benar
50	mengkeke	keke	keke	Benar
51	kiison	kiis	kiis	Benar
52	kioron	kior	kior	Benar
53	kokirag	kirag	kirag	Benar
54	kolonon	kolon	kolon	Benar
55	mongkolon	kolon	kolon	Benar
56	nikolon	kolon	kolon	Benar
57	pongkorowian	korowi	korowi	Benar
58	kuneqon	kuneq	kuneq	Benar
59	mengeleeloq	leeloq	leeloq	Benar
60	nileengket	leengket	leengket	Benar
61	malesiq	lesiq	lesiq	Benar
62	mengelewoti	lewot	lewot	Benar
63	limangon	limang	limang	Benar
64	linsonon	linson	linson	Benar
65	moloq	loeq	loeq	Benar
66	mongololoikon	loloikon	loloikon	Benar
67	mongololoq	loloq	loloq	Benar
68	moloop	loop	loop	Benar
69	poloopan	loop	loop	Benar
70	mongolopol	lopol	lopol	Benar
71	pongolopol	lopol	lopol	Benar
72	molusaq	lusaq	lusaq	Benar
73	ngaanon	ngaan	ngaan	Benar
74	pangaan	ngaan	ngaan	Benar



### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Berimbuhan	Hasil Validator	Hasil Algoritma	Hasil
75	manggeo	nggeo	nggeo	Benar
76	monguraq	nguraq	nguraq	Benar
77	paalekon	paale	paale	Benar
78	pakeon	pake	pake	Benar
79	paoqon	paoq	paoq	Benar
80	parentaon	parenta	parenta	Benar
81	paresaon	paresa	paresa	Benar
82	niparot	parot	parot	Benar
83	pesoon	peso	peso	Benar
84	kopii	pii	kopi	Salah
85	piison	piis	piis	Benar
86	pisokon	pisok	pisok	Benar
87	mapolos	polos	polos	Benar
88	popokon	popok	popok	Benar
89	niposuuq	posuuq	posuuq	Benar
90	puarion	puari	puari	Benar
91	niputuskon	putus	putus	Benar
92	raeqan	raeq	raeq	Benar
93	mangarakati	rakat	rakat	Benar
94	mangarakop	rakop	rakop	Benar
95	pangarakop	rakop	rakop	Benar
96	kabantangan	bantang	bantang	Benar
97	nobebeq	bebeq	bebeq	Benar
98	noboiti	boit	boit	Benar
99	kakaan	kaan	kaan	Benar
100	nokarei	karei	karei	Benar
101	molikison	likison	likison	Benar
102	noliliqon	liliq	liliq	Benar
103	palimangon	limang	limang	Benar
104	loloikonon	loloikon	loloikon	Benar
105	nomaeq	maeq	maeq	Benar
106	nomaeqan	maeq	maeq	Benar
107	nakaton	nakat	nakat	Benar
108	nanauq	nauq	nauq	Benar
109	sanondom	nondom	nondom	Benar
110	kooloa	oloa	oloa	Benar
111	peperai	perai	pera	Salah
112	peperaion	perai	perai	Benar
113	posuuqon	posuuq	posuuq	Benar
114	kopudat	pudat	pudat	Benar
115	mangarapatkon	rapat	rapat	Benar
116	norapat	rapat	rapat	Benar
117	rarangon	rarang	rarang	Benar
118	mengerembas	rembas	rembas	Benar

### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Berimbuhan	Hasil Validator	Hasil Algoritma	Hasil
119	porobuan	robu	robu	Benar
120	norude	rude	rude	Benar
121	koruntun	runtun	runtun	Benar
122	mansagia	sagia	sagia	Benar
123	sakanon	sakan	sakan	Benar
124	mansambai	samba	samba	Benar
125	sampaqon	sampaq	sampaq	Benar
126	sangkilan	sangkil	sangkil	Benar
127	mansarak	sarak	sarak	Benar
128	saruion	sarui	sarui	Benar
129	masasom	sasom	sasom	Benar
130	saweqan	saweq	saweq	Benar
131	kesenian	seni	seni	Benar
132	kosian	sian	sian	Benar
133	sidaon	sida	sida	Benar
134	sinangaon	sinanga	sinanga	Benar
135	monsoani	soan	soan	Benar
136	soputon	soput	soput	Benar
137	monsoput	soput	soput	Benar
138	suntikon	suntik	suntik	Benar
139	nosuoq	suq	suq	Benar
140	surion	suri	suri	Benar
141	suulion	suuli	suuli	Benar
142	notaka	taka	taka	Benar
143	takalan	takal	takal	Benar
144	mantaqan	taqan	taqan	Benar
145	pantaqanan	taqan	taqan	Benar
146	tanggungan	tanggung	tanggung	Benar
147	taraion	tarai	tarai	Benar
148	pantaraian	tarai	tarai	Benar
149	taraqon	taraq	taraq	Benar
150	mantarima	tarima	tarima	Benar
151	taringon	taring	taring	Benar
152	mantatapi	tatapi	tatapi	Benar
153	taunan	taun	taun	Benar
154	sataun	taun	taun	Benar
155	mantausi	tausi	tausi	Benar
156	tiason	tias	tias	Benar
157	motias	tias	tias	Benar
158	timbangan	timbang	timbang	Benar
159	timbangon	timbang	timbang	Benar
160	nitoonkon	toon	toon	Benar
161	tooyukon	tooyuk	tooyuk	Benar
162	montoooyuk	tooyuk	tooyuk	Benar

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Berimbuhan	Hasil Validator	Hasil Algoritma	Hasil
163	pontopongian	topongi	topongi	Benar
164	montotoboq	totoboq	totoboq	Benar
165	nitotoboqi	totoboq	totoboq	Benar
166	totoboqon	totoboq	totoboq	Benar
167	motuqa	tuqa	tuqa	Benar
168	katuqu	tuqu	tuqu	Benar
169	tulungi	tulung	tulung	Benar
170	nitutuk	tutuk	tutuk	Benar
171	umumkon	umum	umum	Benar
172	mongupa	upa	upa	Benar
173	upaon	upa	upa	Benar
174	mangawawa	wawa	wawa	Benar
175	wawaon	wawa	wawa	Benar
176	mangawawau	wawau	wawau	Benar
177	mangawawau	wawau	wawau	Benar
178	pangawawau	wawau	wawau	Benar
179	wawauon	wawau	wawau	Benar
180	weeron	weer	weer	Benar
181	mongowoot	woot	woot	Benar
182	niwoot	woot	woot	Benar
183	wooton	woot	woot	Benar
184	wawauon	wawau	wawau	Benar
185	memboit	boit	boit	Benar
186	mempalau	palau	palau	Benar
187	mempakana	pakana	pakana	Benar
188	membontulkon	bontul	bontul	Benar
189	membontili	bontili	bontili	Benar
190	montanati	tanati	tanati	Benar
191	mondamaikon	damai	damai	Benar
192	monsangkul	sangkul	sangkul	Benar
193	monsampur	sampur	sampur	Benar
194	monjanji	janji	janji	Benar
195	monjaga	jaga	jaga	Benar
196	mongkeke	keke	keke	Benar
197	monggampangkon	gampang	gampang	Benar
198	monghansurkon	hansur	hansur	Benar
199	monulei	ulei	ule	Salah
200	monampedas	ampedas	ampedas	Benar
201	moninum	inum	inum	Benar
202	mangalayak	layak	layak	Benar
203	mangaleot	leot	leot	Benar
204	manganakat	nakat	nakat	Benar
205	mangarawat	rawat	rawat	Benar
206	pongkoyo	koyo	koyo	Benar



#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Berimbuhan	Hasil Validator	Hasil Algoritma	Hasil
207	pongoloe	loe	loe	Benar
208	pangaruruki	ruruki	ruruk	Salah
209	pengeleotan	leot	leot	Benar
210	pengekekan	keke	keke	Benar
211	sapon	sap	sap	Benar
212	tulunon	tulun	tulun	Benar
213	bontulon	bontul	bontul	Benar
214	kahansuran	hansur	hansur	Benar
215	kaidekan	idek	idek	Benar
216	kabosian	bosi	bosi	Benar
217	pankaamatan	kaamat	kaamat	Benar
218	panalusan	alus	alus	Benar
219	salipu	lipu	lipu	Benar
220	sakarung	karung	karung	Benar
221	sarombongan	rombongan	rombongan	Benar
222	saember	ember	ember	Benar
223	mongkot	kot	kot	Benar
224	mongguas	guas	guas	Benar
225	mongulei	ulei	ule	Salah
226	mongkait	kait	kait	Benar
227	mongorurukikon	ruruki	ruruki	Benar
228	mongguaskon	guas	guas	Benar
229	mongkotkon	kot	kot	Benar
230	mongatorkon	ator	ator	Benar
231	menggugurkon	gugur	gugur	Benar
232	montusikon	tusi	tusi	Benar
233	sampurkon	sampur	sampur	Benar
234	sobakon	soba	soba	Benar
235	tusikon	tusi	tusi	Benar
236	loekon	loe	loe	Benar
237	inumkon	inum	inum	Benar
238	niator	ator	ator	Benar
239	nigugur	gugur	gugur	Benar
240	niinti	inti	inti	Benar
241	nikoyo	koyo	koyo	Benar
242	nilayak	layak	layak	Benar
243	monggambar	gambar	gambar	Benar
244	mombose	bose	bose	Benar
245	monsonto	sonto	sonto	Benar
246	mombuanig	buaniq	buaniq	Benar
247	monpaat	paat	paat	Benar
248	mongampudui	ampudu	ampudu	Benar
249	mongalalom	alalom	alalom	Benar
250	mongalusi	alus	alus	Benar

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Berimbuhan	Hasil Validator	Hasil Algoritma	Hasil
251	monghansur	hansur	hansur	Benar
252	mombalakiq	balakiq	balakiq	Benar
253	mansababkon	sabab	sabab	Benar
254	mansampekon	sampe	sampe	Benar
255	manglawan	lawan	lawan	Benar
256	mambayar	bayar	bayar	Benar
257	pambayaran	bayar	bayar	Benar
258	sarakon	sarak	sarak	Benar
259	bayaron	bayar	bayar	Benar
260	panjaraon	panjara	panjara	Benar
261	panjarakon	panjara	panjara	Benar
262	mangkaraniq	karaniq	karaniq	Benar
263	mantangkis	tangkis	tangkis	Benar
264	nipimpin	pimpin	pimpin	Benar
265	mensanggiran	sanggiran	sanggiran	Benar
266	nirakop	rakop	rakop	Benar
267	marawat	rawat	rawat	Benar
268	mangawawu	wawu	wawu	Benar
269	mangarat	arat	arat	Benar
270	mansarui	sarui	sarui	Benar
271	mentandakon	tanda	tanda	Benar
272	mengkeook	keook	keook	Benar
273	monsosop	sosop	sosop	Benar
274	mantarai	tarai	tara	Salah
275	mampatei	pate	pate	Benar
276	mapanas	panas	panas	Benar
277	pae-pae	pae	pae	Benar
278	loka-loka	loka	loka	Benar
279	siok-siok	siok	siok	Benar
280	palo-palo	palo	palo	Benar
281	nantin-nantin	nantin	nantin	Benar
282	rana-rana	rana	rana	Benar
283	sampur-sampur	sampur	sampur	Benar
284	sangkul-sangkul	sangkul	sangkul	Benar
285	montulun-tulun	tulun	tulun	Benar
286	monsoba-soba	soba	soba	Benar
287	bontul-bontul	bontul	bontul	Benar
288	ule-ule	ule	ule	Benar
289	kolon-kolon	kolon	kolon	Benar
290	kodi-kodi	kodi	kodi	Benar
291	olola-olola	olola	olola	Benar
292	donan-donan	donan	donan	Benar
293	kolon	kolon	kolon	Benar
294	mokolon	kolon	kolon	Benar

#### Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Kata Berimbuhan	Hasil Validator	Hasil Algoritma	Hasil
295	atorkon	ator	ator	Benar
296	mambarang	barang	barang	Benar
297	bela	bela	bela	Benar
298	bobok	bobok	bobok	Benar
299	momboit	boit	boit	Benar
300	momboloskon	bolos	bolos	Benar
301	bolos	bolos	bolos	Benar
302	mandaga	daga	daga	Benar
303	mondoloqi	doloq	doloq	Benar
304	mongimput	imput	imput	Benar
305	monginum	inum	inum	Benar
306	kambur	kambur	kambur	Benar
307	mankaris	karis	karis	Benar
308	mankaris	karis	karis	Benar
309	mapeso	peso	peso	Benar
310	mapoloskon	polos	polos	Benar



#### Hak Cipta Dilindungi Undang-Undang

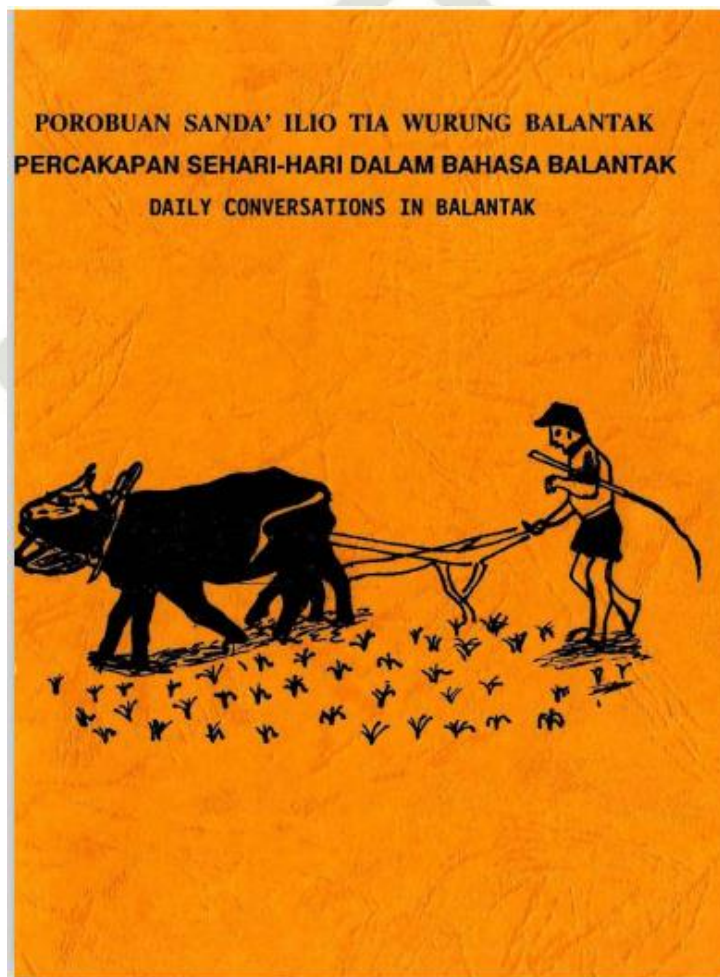
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LAMPIRAN D

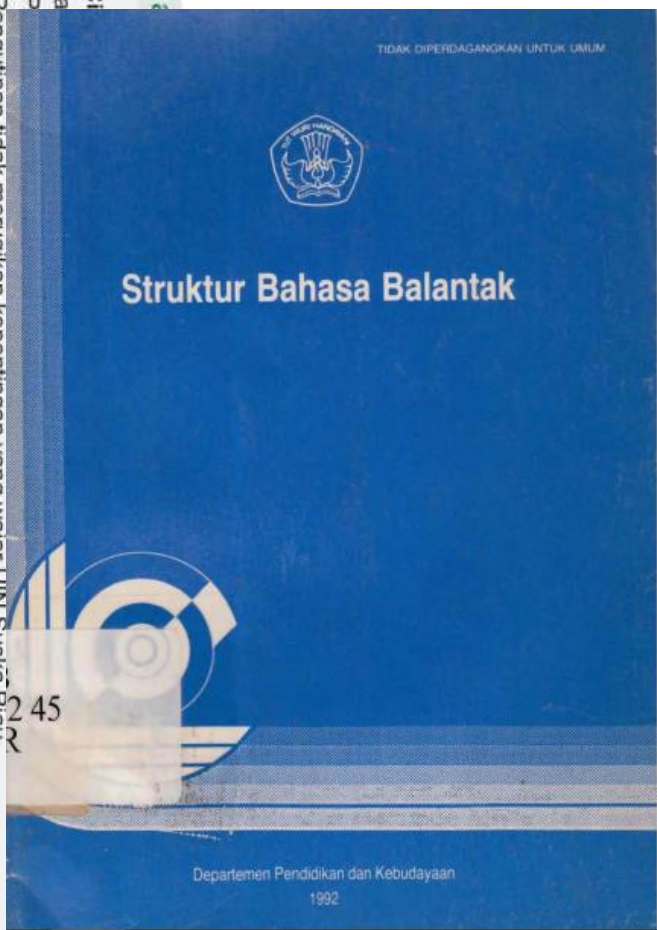
### DATA PENELITIAN

Transkrip ini merupakan data penelitian yang diperoleh dalam membangun algoritma *stemming* teks pada bahasa Balantak.

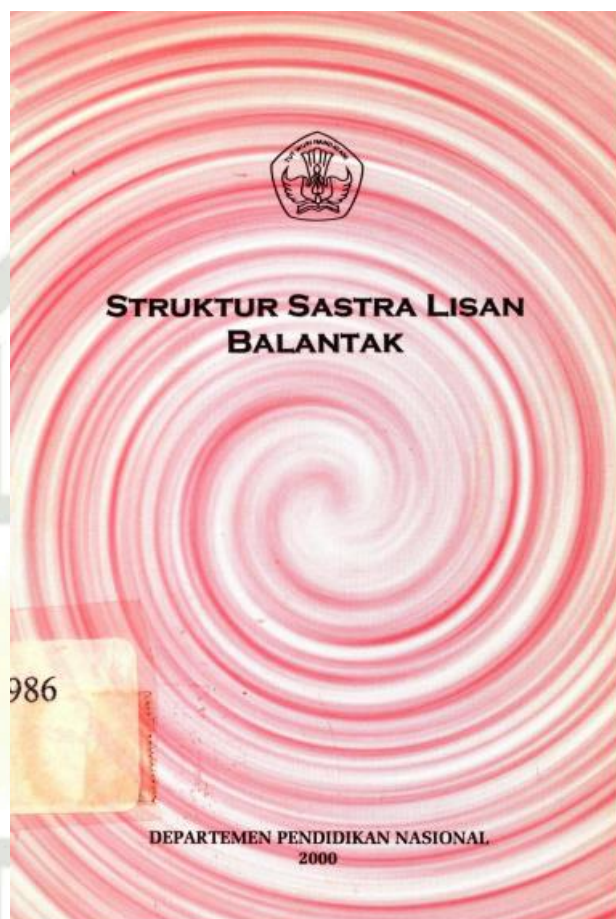
#### A. Buku Percakapan Sehari-hari dalam Bahasa Balantak



B. Buku Struktur Bahasa Balantak



C. Buku Struktur Sastra Lisan Balantak



## DAFTAR RIWAYAT HIDUP

Nama Lengkap : Surya Atmaja  
Tempat/Tanggal Lahir : Lahang Hulu / 15 Oktober 1995  
Nama Ayah : Antoni  
Nama Ibu : Rosminar  
Anak ke : 5  
Jumlah Sdr : 5  
Alamat : Jl. Handayani GG. Mawar Marpoyan Damai Pekanbaru  
E-mail : Surya.atmaja@students.uin-suska.ac.id

## PENDIDIKAN

- ✓ Tahun 2001-2007 : SD Negeri 009 Gaung
- ✓ Tahun 2007-2010 : SMP Negeri 1 Gaung
- ✓ Tahun 2010-2013 : SMA Negeri 1 Gaung
- ✓ Tahun 2013-2020 : Universitas Islam Negeri Sultan Syarif Kasim Riau.

Jurusan Teknik Informatika.

UIN SUSKA RIAU

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.